INRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél: 954 90 20

Rapports de Recherche

N° 170

GRAPHTHEORETICAL CLASSIFICATION AND GENERATION OF JD'S BY MEANS OF BASES FOR MVD-STRUCTURES

Yoshito HANATANI

Novembre 1982

GRAPHTHEORETICAL CLASSIFICATION AND

GENERATION OF JD'S BY MEANS OF

BASES FOR MVD-STRUCTURES

Yoshito HANATANI

ABSTRACT

Starting from characterizations of a set of MVD's implied by a single JD, we propose a graph-theoretical approach for classifying and generating JD's. The basic ingredient of our approach is the concept of basis for MVD-structures, presented in an earlier paper.

RESUME

Partant des caractérisations d'un ensemble de dépendances multi-valuées impliqué par une seule dépendance de jointure, on propose une méthode utilisant la théorie des graphes pour classifier et générer des dépendances de jointure. L'outil principal de notre méthode est la représentation en termes de base d'un ensemble de dépendances multi-valuées, que nous avons introduite dans un papier précédent.

INTRODUCTION

Let U be a finite set of attributes. Let MVD(U) and JD(U) be the set of all the MVD's (multi-valued dependencies) and that of all the JD's (join dependencies) on U, respectively. Formally they are sets of expressions as follows :

$$MVD(U) =_{def} \{X \twoheadrightarrow Y \mid X,Y \subseteq U\}$$

$$JD(U) =_{def} \{\bowtie \underline{X} \mid \underline{X} \subseteq P(U) - \{\emptyset\}, \bigcup \underline{X} = U\}.$$

[BFMY] studied several characterizations of subsets M of MVD(U) for which there exists a JD $j \in JD(U)$ such that $M =_{MVD} \{j\}$; the MDV-equality '$=_{MVD}$' is an equivalence relation between sets C_1, C_2 of integrity constraints of any sort, defined by :

$$C_1 =_{MVD} C_2 \Longleftrightarrow_{def} \forall\, d \in MVD(U)\ (C_1 \models d \Longleftrightarrow C_2 \models d).$$

In this definition, '$\models$' is the semantical implication in the usual sense, i.e. $C \models d \Longleftrightarrow$ 'for any relation R on U, if $R \mid\!\sim C$ (R satisfies C) then $R \mid\!\sim d$'.

For every $M \subseteq MVD(U)$, let us define a subset of JD(U) as follows :

$$jds(M) =_{def} \{j \in JD(U) \mid \{j\} =_{MVD} M\}$$

Then the above mentioned characterizations are those of subsets $M \subseteq MVD(U)$ such that $jds(M) \neq \emptyset$.

There is a notion of basis for every subset M of MVD(U) (see [H]). A basis $\mathcal{B}$ is a family of subsets of U and it characterizes an equivalence class of subsets of MVD(U) with respect to '$=_{MVD}$'.

In the present paper we look for conditions so that $jds(M) \neq \emptyset$, in terms of a basis of M. Also, for ways to generate all the elements of jds(M) via a basis $\mathcal{B}$ of M. That is, we want to know the graphtheoretical relation between the elements of jds(M) and a basis $\mathcal{B}$ of M. We want to know even, when a JD j is given, how to obtain the other elements of jds(M) which contains j, and how to obtain $\mathcal{B}$ which is a basis of M such that $\{j\} =_{MVD} M$.

The notion of basis relies on the notion of agreement relation. The agreement relation 'B agrees with d' is a relation between a subset B of U and an element d of MVD(U) and it is defined as follows :

$$B \rightsquigarrow X \twoheadrightarrow Y \iff_{def} (X \subseteq B \Longrightarrow Y \subseteq B \vee (U - Y) \subseteq B).$$

It is extended to the set levels by the universal quantification as follows :

$$\mathcal{B} \rightsquigarrow d \iff_{def} \forall B \in \mathcal{B} \, (B \rightsquigarrow d),$$

$$B \rightsquigarrow M \iff_{def} \forall d \in M \, (B \rightsquigarrow d),$$

$$\mathcal{B} \rightsquigarrow M \iff_{def} \forall B \in \mathcal{B} \, \forall d \in M \, (B \rightsquigarrow d).$$

Using this relation, we define a pair of set functions as follows :

$$\mathcal{P}(\mathrm{MVD}(U)) \underset{\mathrm{BASE}^{-1} < U, \mathrm{MVD} >}{\overset{\mathrm{BASE} < U, \mathrm{MVD} >}{\rightleftarrows}} \mathcal{P}(\mathcal{P}(U))$$

$$\mathrm{BASE} < U, \mathrm{MVD} > (M) =_{def} \{B \in \mathcal{P}(U) \mid B \rightsquigarrow M\},$$

$$\mathrm{BASE}^{-1} < U, \mathrm{MVD} > (\mathcal{B}) =_{def} \{d \in \mathrm{MVD}(U) \mid \mathcal{B} \rightsquigarrow d\}.$$

We abbreviate them as BASE(M) and $\mathrm{BASE}^{-1}(\mathcal{B})$. They are order-inverting functions w.r.t. to inclusion, satisfying the following relations :

(B1) $M \subseteq \mathrm{BASE}^{-1} \circ \mathrm{BASE}\ (M)$ for any $M \subseteq \mathrm{MVD}(U)$

(B2) $\mathcal{B} \subseteq \mathrm{BASE} \bullet \mathrm{BASE}^{-1}\ (\mathcal{B})$ for any $\mathcal{B} \subseteq \mathcal{P}(U)$

It is proved in [H] that

(B3) $\mathrm{BASE}(M) \supset\!\sim d \iff M \models d$

for any $M \subseteq \mathrm{MVD}(U)$ and for any $d \in \mathrm{MVD}(U)$.

This allows us to define the following notion :

$\mathcal{B}$ is a basis of M

$\iff_{\mathrm{def}} \forall\, d \in \mathrm{MVD}(U)\ (\mathcal{B} \supset\!\sim d \iff \mathrm{BASE}\ (M) \supset\!\sim d)$,

where $\mathcal{B}$ is any subset of $\mathcal{P}(U)$ and M is any subset of MVD(U).

It should be noted that :

$\mathcal{B}$ is a basis of M

$\iff \forall\, B \in \mathcal{P}(U)\ (B \supset\!\sim \mathrm{BASE}^{-1}(\mathcal{B}) \iff B \supset\!\sim M)$,

and that BASE(M) is the unique maximal basis of M.

In what follows, we present two results :

A. characterizations of subsets $\mathcal{B}$ of $\mathcal{P}(U)$ such that

$\exists M \subseteq \mathrm{MVD}(U)\ (\mathcal{B}$ is a basis of $M \wedge \mathrm{jds}(M) \neq \emptyset)$,

B. a graphtheoretical approach to the relation between jds(M) and some subset $\mathcal{B}$ of $\mathcal{P}(U)$ wich is a basis of M, and to the relation between the elements of jds(M).

PART A. CHARACTERIZATIONS

Our aim here is to characterize those subsets $\mathcal{B}$ of $\mathcal{P}(U)$, that are a basis of some subset M of MVD(U), such that $\mathrm{jds}(M) \neq \emptyset$.

The condition $\mathrm{jds}(M) \neq \emptyset$ is shown in [BFMY] to be equivalent to the intersection property and also to the orthogonality property on M. These properties can be stated as follows.

Int (M) :

$$\forall X,Y,Z \subseteq U \; (Z \subseteq U - X \cup Y \wedge M \models X \twoheadrightarrow Z \wedge M \models Y \twoheadrightarrow Z \Rightarrow M \models X \cap Y \twoheadrightarrow Z),$$

Orth (M) :

$$\forall X,Y \subseteq U \; (X \cap Y = \emptyset \wedge \forall x \in X \, \forall y \in Y \, (M \models U - \{x,y\} \twoheadrightarrow \{x\}) \Rightarrow M \models U - X \cup Y \twoheadrightarrow X).$$

For any subset $\mathcal{B}$ of $\mathcal{P}(U)$, let us consider the following condition.

Int ($\mathcal{B}$) :

$$\forall X,Y,Z \subseteq U \; (Z \subseteq U - X \cup Y \wedge \mathcal{B} \mathrel{\supset\!\sim} X \twoheadrightarrow Z \wedge \mathcal{B} \mathrel{\supset\!\sim} Y \twoheadrightarrow Z \Rightarrow \mathcal{B} \mathrel{\supset\!\sim} X \cap Y \twoheadrightarrow Z).$$

Then we see immediately that

$$\mathrm{jds}(M) \neq \emptyset \iff \mathrm{Int}(\mathrm{BASE}(M)),$$

because $\mathrm{Int}(M) \iff \mathrm{Int}(\mathrm{BASE}(M))$, by the equivalence

(B3) : $\forall M \subseteq \mathrm{MVD}(U) \; \forall d \in \mathrm{MVD}(U) \; (\mathrm{BASE}(M) \mathrel{\supset\!\sim} d \iff M \models d)$.

Now let $\mathcal{B}$ be a basis of M, i.e.

$$\forall d \in \mathrm{MVD}(U) \; (\mathcal{B} \mathrel{\supset\!\sim} d \iff \mathrm{BASE}(M) \mathrel{\supset\!\sim} d).$$

Then :

$$\mathrm{jds}(M) \neq \emptyset \iff \mathrm{Int}\ (\mathcal{B}),$$

because we then trivially have : Int (BASE(M)) $\iff$ Int ($\mathcal{B}$). Thus we have shown that :

(1) B is a basis of M $\Rightarrow$ (jds(M) $\neq \emptyset \Leftrightarrow$ Int(B)).

We want to show that the condition Int (B) characterizes B such that

$\exists M \subseteq MVD(U)$ (B is a basis of M $\wedge$ jds(M) $\neq \emptyset$).

Lemma 1

The following two conditions are equivalent.

1) $\exists M \subseteq MVD(U)$ (B is a basis of M $\wedge$ jds(M) $\neq \emptyset$)

2) $\forall M \subseteq MVD(U)$ (B is a basis of M $\Rightarrow$ jds(M) $\neq \emptyset$).

Proof

1) $\Rightarrow$ 2) : this is a consequence of the fact :

B is a basis of M and also a basis of M' $\Rightarrow$ $M =_{MVD} M'$,

for any $M, M' \subseteq MVD(U)$, by the equivalence (B3).

2) $\Rightarrow$ 1) : this is immediate because B is a basis of $BASE^{-1}(B)$. □

Theorem 1 (a characterization)

For any $B \subseteq P(U)$, the condition Int(B) is equivalent to the condition :

$\exists M \subseteq MVD(U)$ (B is a basis of M $\wedge$ jds(M) $\neq \emptyset$).

Proof

Immediate by (1) and Lemma 1. □

Now let us consider the following condition on any subset B of $P(U)$, that we shall call the minus-two-completeness of B ;

Min 2 (B) :

$$\forall X,Y \subseteq U \; (X \cap Y = \emptyset \wedge X \neq \emptyset \wedge Y \neq \emptyset \wedge U - X \cup Y \in B$$
$$\Rightarrow \exists x \in X \; \exists y \in Y \; (U-\{x,y\} \in B)).$$

Then we can state that

(2) $\mathrm{jds}(M) \neq \emptyset \iff \mathrm{Min}\ 2(\mathrm{BASE}(M))$,

for any $M \subseteq \mathrm{MVD}(U)$,

because $\mathrm{Orth}(M) \Leftrightarrow \mathrm{Min2}(\mathrm{BASE}(M))$, as we see it easily. When B is a basis of M,

$$\mathrm{BASE}(M) = \mathrm{BASE} \circ \mathrm{BASE}^{-1}(B).$$

So we have the following :

(2') B is a basis of $M \Rightarrow (\mathrm{jds}(M) \neq \emptyset \Leftrightarrow \mathrm{Min}\ 2(\mathrm{BASE} \circ \mathrm{BASE}^{-1}\ (B)))$.

Thus we have got another characterization :

THEOREM 2 (a characterization)

For any $B \subseteq P(U)$, the condition $\mathrm{Min2}(\mathrm{BASE} \circ \mathrm{BASE}^{-1}(B))$ is equivalent to the condition :

$$\exists M \subseteq \mathrm{MVD}(U) \; (B \text{ is a basis of } M \wedge \mathrm{jds}\ (M) \neq \emptyset).$$

Proof

Immediate by (2') and Lemma 1. □

We shall introduce one more characterization. Let B be any subset of $P(U)$. The minus two component of B, denoted by B_{-2}, is the subset of B, defined as follows :

$$B_{-2} =_{\mathrm{def}} \{B \in B \mid \mathrm{card}\ B = \mathrm{card}\ U - 2\}.$$

Lemma 2

Let B and B' be any subsets of $P(U)$. Then :

$$\mathrm{BASE}^{-1}(\mathcal{B}) \subseteq \mathrm{BASE}^{-1}(\mathcal{B}') \Longrightarrow \mathcal{B}'_{-2} \subseteq \mathcal{B}_{-2} .$$

Proof

Let B' be any element of $\mathcal{B}'_{-2}$. Let $\{x,y\} = U - B'$. Then $B' \twoheadrightarrow \{x\} \notin (\mathrm{BASE}^{-1}(\mathcal{B}'))$. So by the hypothesis , $B' \twoheadrightarrow \{x\} \notin \mathrm{BASE}^{-1}(\mathcal{B})$. This means that there is some B in $\mathcal{B}$ such that $B' \subseteq B$, $\{x\} \not\subseteq B$ and $\{y\} \subseteq B$. But such a B must be equal to B'. That is, B' is in $\mathcal{B}$, therefore in $\mathcal{B}_{-2}$. □

Corollary 1

For any subset $\mathcal{B}$ of $P(U)$,

$$(\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}))_{-2} \subseteq \mathcal{B}_{-2} .$$

Proof

Because $\mathrm{BASE}^{-1}(\mathcal{B}) \subseteq \mathrm{BASE}^{-1} \circ \mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B})$, by the property (B2). □

Corolly 2 (Uniqueness)

For any subsets $\mathcal{B}$ and $\mathcal{B}'$ of $P(U)_{-2}$,

$$\exists M \subseteq \mathrm{MVD}(U)\ (\mathcal{B} \text{ is a basis of } M \text{ and } \mathcal{B}' \text{ is a base of } M) \Longleftrightarrow \mathcal{B} = \mathcal{B}'$$ □

Theorem 3 (a characterization)

For any subset $\mathcal{B}$ of $P(\mathbf{U})$, the condition Min 2 Rep($\mathcal{B}$) (the minus-two-representability) : $\mathrm{BASE}^{-1}(\mathcal{B}_{-2}) \subseteq \mathrm{BASE}^{-1}(\mathcal{B})$, is equivalent to the condition :

$$\exists M \subseteq \mathrm{MVD}(U)\ (\mathcal{B} \text{ is a basis of } M \wedge \mathrm{jds}(M) \neq \emptyset).$$

Proof

It suffices to show the following two implications :

1) Min 2 $(\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B})) \Longrightarrow$ Min 2 Rep$(\mathcal{B})$,

2) Min 2 Rep$(\mathcal{B}) \Longrightarrow$ Int $(\mathcal{B})$.

<u>Proof of 1)</u> Let d be any element of MVD(U) such that $d \notin \mathrm{BASE}^{-1}(\mathcal{B})$. We want to show that $d \notin \mathrm{BASE}^{-1}(\mathcal{B}_{-2})$. The assumption $d \notin \mathrm{BASE}^{-1}(\mathcal{B})$ implies $\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}) \not\rightsquigarrow d$. So by Min 2 $(\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}))$, we have $(\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}))_{-2} \not\rightsquigarrow d$. It follows that $\mathcal{B}_{-2} \not\rightsquigarrow d$ by corollary1.That is, $d \notin \mathrm{BASE}^{-1}(\mathcal{B}_{-2})$.

<u>Proof of 2)</u> Assume that $\mathcal{B} \not\rightsquigarrow X \cap Y \twoheadrightarrow Z$ with $Z \subseteq U - X \cup Y$. Then, by Min 2 Rep$(\mathcal{B})$, we have $\mathcal{B}_{-2} \not\rightsquigarrow X \cap Y \twoheadrightarrow Z$. Let B be any element of $\mathcal{B}_{-2}$ such that $B \not\rightsquigarrow X \cap Y \twoheadrightarrow Z$. We shall show that B satisfies either $B \not\rightsquigarrow X \twoheadrightarrow Z$ or $B \not\rightsquigarrow Y \twoheadrightarrow Z$, which is sufficient to conclude that either $\mathcal{B} \not\rightsquigarrow X \twoheadrightarrow Z$ or $\mathcal{B} \not\rightsquigarrow Y \twoheadrightarrow Z$. Let $\{x,y\} = U - B$ with $x \in Z$ and $y \in U - Z - X \cap Y$. Then $x \notin X \cup Y$ by $Z \subseteq U - X \cup Y$. Thus we have shown that either $B \not\rightsquigarrow X \twoheadrightarrow Z$ or $B \not\rightsquigarrow Y \twoheadrightarrow Z$. □

<u>Corollary 3</u> (Principal application)

For any subset M of MVD(U),

$\mathrm{jds}(M) \neq \emptyset \Longleftrightarrow \exists \mathcal{B} \subseteq \mathcal{P}(U)_{-2}$ ($\mathcal{B}$ is a basis of M).

<u>Proof</u>

$\Longrightarrow$) Assume $\mathrm{jds}(M) \neq \emptyset$. Let $\mathcal{B}' = \mathrm{BASE}(M)$. Then $\mathcal{B}'$ is a basis of M. So, by Theorem 3,

$$\forall d \in \mathrm{MVD}(U)(\mathcal{B}' \rightsquigarrow d \Longleftrightarrow \mathcal{B}'_{-2} \rightsquigarrow d).$$

But the fact that $\mathcal{B}'$ is a basis of M, means that

$$\forall d \in \mathrm{MVD}(U)(\mathcal{B}' \rightsquigarrow d \Longleftrightarrow \mathrm{BASE}(M) \rightsquigarrow d).$$

We have therefore :

$$\forall\, d \in \mathrm{MVD(U)}\ (B'_{-2} \supset\!\!\sim d \Leftrightarrow \mathrm{BASE(M)} \supset\!\!\sim d),$$

that is, B'_{-2} is a basis of M. Naturally, $B'_{-2} \subseteq P(\mathrm{U})_{-2}$.

$\Leftarrow$) Let $B \subseteq P(\mathrm{U})_{-2}$. Then $B_{-2} = B$, therefore we have Min 2 Rep (B). But this is equivalent to $\forall\, \mathrm{M}(B$ is a basis of $\mathrm{M} \Rightarrow \mathrm{jds(M)} \neq \emptyset)$ by Theorem 3 and Lemma 1. Now, what we wanted to prove has become clear. □

The above results, in particular Corollaries 2 and 3, show that there is a bijective correspondence between the quotient set $\mathrm{JD(U)} / =_{\mathrm{MVD}}$ and $P(P(\mathrm{U})_{-2})$, the power set of $P(\mathrm{U})_{-2}$. In fact, any equivalence class $[j]_{\mathrm{MVD}}$ which is an element of $\mathrm{JD(U)} / =_{\mathrm{MVD}}$, can be represented as jds(M) with some subset M of MVD(U). So, by Corollary 3, it can be represented as $\mathrm{jds}(\mathrm{BASE}^{-1}(B))$ with $B \subseteq P(\mathrm{U})_{-2}$, because $\mathrm{jds(M)} = \mathrm{jds}(\mathrm{BASE}^{-1} \circ \mathrm{BASE(M)})$ by (B3) and the definition of jds(M). On the other hand, it is clear that any $\mathrm{jds}(\mathrm{BASE}^{-1}(B))$ is an element of $\mathrm{JD(U)} / =_{\mathrm{MVD}}$. It remains to show the following :

<u>Theorem 4</u>

$$B = B' \Leftrightarrow \mathrm{jds}(\mathrm{BASE}^{-1}(B)) = \mathrm{jds}(\mathrm{BASE}^{-1}(B')),$$

for any B, $B' \subseteq P(\mathrm{U})_{-2}$.

<u>Proof</u>

$\Rightarrow$) Trivial.

$\Leftarrow$) Assume that $\mathrm{jds}(\mathrm{BASE}^{-1}(B))$ and $\mathrm{jds}(\mathrm{BASE}^{-1}(B))$ have a common element. Then, by transitivity, $\mathrm{BASE}^{-1}(B) =_{\mathrm{MVD}} \mathrm{BASE}^{-1}(B')$. It remains only to show that :

$$(\ast)\ \mathrm{BASE}^{-1}(\mathcal{B}) =_{\mathrm{MVD}} \mathrm{BASE}^{-1}(\mathcal{B}') \Longrightarrow \mathrm{BASE}^{-1}(\mathcal{B}) = \mathrm{BASE}^{-1}(\mathcal{B}'),$$

because it then follows that $\mathcal{B} = \mathcal{B}'$, by Corollary 2.

Proof of $(\ast)$: we must show that for any $\mathcal{B} \subseteq \mathcal{P}(\mathrm{U})$,

$$\{d \in \mathrm{MVD} \mid \mathrm{BASE}^{-1}(\mathcal{B}) \models d\} = \mathrm{BASE}^{-1}(\mathcal{B}).$$

It suffices to show that $\mathrm{BASE}^{-1}(\mathcal{B}) \models d \Longrightarrow \mathcal{B} \supset\!\sim d$.

The premise is equivalent to $\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}) \supset\!\sim d$, by (B3).
And $\mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B}) \supset\!\sim d \Rightarrow \mathcal{B} \supset\!\sim d$, is evident, when we note the property (B2), i.e. $\mathcal{B} \subseteq \mathrm{BASE} \circ \mathrm{BASE}^{-1}(\mathcal{B})$. This completes the proof of $(\ast)$.

□

PART B. GRAPHTHEORETICAL APPROACH

[FMU] shows how to get from a given JD j, the set MVD(j) of all the MVD's implied by j. Here, we want to study how to generate the set jds(M) from a given subset $\mathrm{M} \subseteq \mathrm{MVD(U)}$.

We have seen in Part A(Corollaries 2 and 3), that we can uniquely determine a subset $\mathcal{B}$ of $\mathcal{P}(\mathrm{U})_{-2}$ as a basis of M, when jds(M) must have at least one element. And for any subset $\mathcal{B}$ of $\mathcal{P}(\mathrm{U})_{-2}$, there is some subsets $\mathrm{M} \subseteq \mathrm{MVD(U)}$, of which $\mathcal{B}$ is a basis and for which $\mathrm{jds(M)} = \mathrm{jds}(\mathrm{BASE}^{-1}(\mathcal{B})) \neq \emptyset$.

So we shall give an algorithm to generate from a given $\mathcal{B} \subseteq \mathcal{P}(\mathrm{U})_{-2}$, all the elements of $\mathrm{jds}(\mathrm{BASE}^{-1}(\mathcal{B}))$.

Let $\mathcal{B}$ be any subset of $\mathcal{P}(\mathrm{U})_{-2}$. We associate to $\mathcal{B}$ a covering of U, which we shall call the inverse set of $\mathcal{B}$, denoted by $\mathrm{I}(\mathcal{B})$. The definition is as follows.

$$\mathrm{I}(\mathcal{B}) =_{\mathrm{def}} \{\mathrm{U} - \mathrm{B} \mid \mathrm{B} \in \mathcal{B}\} \cup \{\{x\} \mid x \in \cap \mathcal{B}\}.$$

We shall show that the JD $\bowtie I(\mathcal{B})$ should be an element of $\mathrm{jds}(\mathrm{BASE}^{-1}(\mathcal{B}))$.

Theorem 5

Let $\mathcal{B}$ be any subset of $\mathcal{P}(U)_{-2}$. Then :

$\forall\, d \in \mathrm{MVD}(U)\ (\bowtie I(\mathcal{B}) \models d \Longleftrightarrow \mathcal{B} \supset\!\!\sim d)$.

Proof

$\Longleftarrow$) Let $\mathcal{B} \supset\!\!\sim X \twoheadrightarrow Y$. Then, by the fact that $\mathcal{B} \subseteq \mathcal{P}(U)_{-2}$, any element B' of $I(\mathcal{B})$ satisfies $B' \cap (Y - X) = \emptyset$ or $B' \cap (U - Y - X) = \emptyset$. And as $I(\mathcal{B})$ is a covering of U, Y must be the union of some connected components of the hypergraph $(U, I(\mathcal{B}))$ with the set of nodes X deleted. It follows by Theorem 3 of FMU that $\bowtie I(\mathcal{B}) \models X \twoheadrightarrow Y$.

$\Longrightarrow$) Let $\mathcal{B} \not\supset\!\!\sim X \twoheadrightarrow Y$. Then there exists $B \in \mathcal{B}$ such that $B \not\supset\!\!\sim X \twoheadrightarrow Y$. Let R be a relation on U, consisting of two tuples f and g such that $f\lceil B\rceil = g\lceil B\rceil$ and for any $x \in U - B$, $f(x) \neq g(x)$. Then it is clear that $R \not\models X \twoheadrightarrow Y$. It suffices to show that $R \models \bowtie I(\mathcal{B})$, in order to conclude that $\bowtie I(\mathcal{B}) \not\models X \twoheadrightarrow Y$. In fact, $R \models \bowtie I(\mathcal{B})$ holds, because any tuple h such that $h\lceil B'\rceil \in R\lceil B'\rceil$ for every $B' \in I(\mathcal{B})$, cannot be other than f or g. □

Corollary 4

Let $\mathcal{B} \subseteq \mathcal{P}(U)_{-2}$ and let $M \subseteq \mathrm{MVD}(U)$. Then :

$\mathcal{B}$ is a basis of $M \Longleftrightarrow \bowtie I(\mathcal{B}) \in \mathrm{jds}(M)$.

Proof It follows from the fact that :

'$\mathcal{B}$ is a basis of M' is equivalent to

$\forall\, d \in \mathrm{MVD}(U)\ (\mathcal{B} \supset\!\!\sim d \Longleftrightarrow M \models d)$,

by definition and by (B3). □

Corollary 5

For any $j \in JD(U)$, there is a unique element of $JD(U)$ of the form $\bowtie I(\mathcal{B})$ with $\mathcal{B} \subseteq \mathcal{P}(U)_{-2}$ such that $j =_{MVD} \bowtie I(\mathcal{B})$

Proof

By Corollaries 2 and 4. □

To know all the elements of $jds(BASE^{-1}(\mathcal{B}))$ is to know the relation '$=_{MVD}$' between the elements of $JD(U)$. Note that $JD(U)$ is the set of all the expressions of the form $\bowtie \underline{X}$ such that $\underline{X}$ is a covering of U. So, our aim is to characterize the relation $\bowtie \underline{X} =_{MVD} \bowtie \underline{Y}$, by means of a relation between two coverings $\underline{X}$ and $\underline{Y}$.

Let $C(U)$ be the set of all the coverings of U, i.e.

$$C(U) =_{def} \{\underline{X} \mid \underline{X} \subseteq \mathcal{P}(U) - \{\emptyset\} \wedge \bigcup \underline{X} = U\}.$$

We want to introduce a reduction relation ':=' between elements of $C(U)$, so that the relation $\bowtie \underline{X} =_{MVD} \bowtie \underline{Y}$ would be equivalent to the relation :

$$\exists\, \underline{Z} \in C(U)\ (\underline{X} := \underline{Z} \text{ and } \underline{Y} := \underline{Z}).$$

Definition

The relation ':=' is a binary relation on $C(U)$. It is defined by unit reductions denoted by '$:=_1$'.

Let $\underline{X}, \underline{Y} \in C(U)$. We say that $\underline{X} := \underline{Y}$ holds, if $\underline{X}$ can be reduced to $\underline{Y}$ by a finite number of successive applications of the unit reductions ; that is, when there is a chain of elements of $C(U)$, say $\underline{X}_0, \underline{X}_1, \ldots, \underline{X}_n$ such that $\underline{X}_0 :=_1 \underline{X}_1, \underline{X}_1 := \underline{X}_2, \ldots, \underline{X}_{n-1} := \underline{X}_n$ and $\underline{X} = \underline{X}_0$, $\underline{Y} = \underline{X}_n$. Consequently, this relation ':=' is reflexive, $\underline{X} := \underline{X}$, and transitive, $X := \underline{Y} \wedge \underline{Y} := \underline{Z} \Rightarrow \underline{X} := \underline{Z}$.

The following three kinds of unit reductions are allowed, where, $\text{Clique}(Y) =_{def} \{Z \mid Z \subseteq Y \wedge \text{card}(Z) = 2\}$.

(r1) $\text{Clique}(Y) \subseteq \underline{X} \Longrightarrow \underline{X} :=_1 \underline{X} \cup \{Y\}$.

(r2) $X \in \underline{X}$ and $\exists X' \in \underline{X} \, (X \subsetneq X') \Rightarrow \underline{X} :=_1 \underline{X} - \{X\}$.

(r3) $\exists X \in \underline{X} \, (Y \subsetneq X) \Longrightarrow \underline{X} :=_1 \underline{X} \cup \{Y\}$.

On the basis of this reduction relation, we define the equality (equivalence relation) as follows.

$$\underline{X} :=: \underline{Y} \Longleftrightarrow_{def} \exists \underline{Z} \in C(U) \; (\underline{X} := \underline{Z} \wedge \underline{Y} := \underline{Z}).$$ □

<u>Lemma 3</u> (Church-Rosser Property)

For any $\underline{X}, \underline{Y}, \underline{Z} \in C(U)$,

$$\underline{X} := \underline{Y} \;\wedge\; \underline{X} := \underline{Z} \Longrightarrow \exists \underline{W} \in C(U) (\underline{Y} := \underline{W} \wedge \underline{Z} := \underline{W})$$

<u>Proof</u>

It is a consequence of the following parallelogram property :

Let $\underline{X} :=_1 \underline{Y}$ by (ri) and $\underline{X} :=_1 \underline{Z}$ by (rj), then there exists $\underline{W}$ such that $\underline{Y} :=_1 \underline{W}$ by (rj) and $\underline{Z} :=_1 \underline{W}$ by (ri), where (ri) and (rj) stand for any of the three rules of unit reduction. □

<u>Corollary 6</u>

The equality :=: is reflexive, symmetric and transitive. □

<u>Lemma 4</u>

For any $\underline{X}, \underline{Y} \in C(U)$,

$$\underline{X} :=_1 \underline{Y} \Longrightarrow \bowtie \underline{X} =_{MVD} \bowtie \underline{Y}.$$

Proof

Every unit reduction is a reduction which does not disturb the connectivity in the corresponding hypergraphs. So, if $\underline{X} :=_1 \underline{Y}$, then the hypergraphs $(U,\underline{X})$, $(U,\underline{Y})$ have the same set of connected components. It follows by Theorem 3 of [FMU] that $\bowtie \underline{X} \models d \Leftrightarrow \bowtie \underline{Y} \models d$ for any $d \in MVD(U)$. □

Corollary 7

For any $\underline{X}$, $\underline{Y} \in C(U)$,

$$\underline{X} :=: \underline{Y} \Longrightarrow \bowtie \underline{X} =_{MVD} \bowtie \underline{Y}.$$ □

It remains to show the inverse. But, since the equality ':=:' is transitive , (Corollary 6), it suffices to find an intermediate element from the condition $\bowtie \underline{X} =_{MVD} \bowtie \underline{Y}$. And we know that we may take as an intermediate element a covering of the form $I(\mathcal{B})$ with $\mathcal{B} \subseteq \mathcal{P}(U)_{-2}$. That is :

Lemma 5

For any $\underline{X} \in C(U)$, we can effectively find an element of $C(U)$ of the form $I(\mathcal{B})$ with $\mathcal{B} \subseteq \mathcal{P}(U)_{-2}$, for which $I(\mathcal{B}) := \underline{X}$ holds ($\mathcal{B}$ may be empty).

Proof

By successive application of (r2) on $\underline{X}$, we can obtain $\underline{Y} \in C(U)$ which is irreducible with respect to (r2). We next replace all $Y \in \underline{Y}$ such that card $(Y) \geq 3$, by the elements of Clique (Y) and we obtain $Z \in C(U)$, satisfying the following two conditions

(1) $\forall Z \in \underline{Z}$ (card$(Z) \leq 2$).

(2) $\forall Z \in \underline{Z} \; \neg \exists Z' \in \underline{Z} \; (Z' \subsetneq Z)$.

$\underline{Z}$ is then of the form $I(B)$ with $B \subseteq P(U)_{-2}$.

It is easy to show that $\underline{Z} := \underline{X}$. In fact, if $\underline{Z} \neq \underline{Y}$, we can obtain $\underline{Y}$ from $\underline{Z}$ by repeated applications of (r1) followed by those of (r2), and $\underline{X}$ from $\underline{Y}$ by repeated applications of (r3). □

Theorem 6

For any $\underline{X}, \underline{Y} \in C(U)$,

$$\bowtie \underline{X} =_{MVD} \bowtie \underline{Y} \iff \underline{X} :=: \underline{Y}.$$

Proof

$\Longleftarrow$) By Corollary 7.

$\Longrightarrow$) By Lemma 5, there exist B and $B' \subseteq P(U)_{-2}$ such that $I(B) := \underline{X}$ and $I(B') := \underline{Y}$. As $I(B) := \underline{X}$ and $I(B') := \underline{Y}$ imply $I(B) :=: \underline{X}$ and $I(B') :=: \underline{Y}$, It follows by Corollary 7, that $\bowtie I(B) =_{MVD} \bowtie \underline{X}$ and $\bowtie I(B') =_{MVD} \bowtie \underline{Y}$.
Now assume that $\bowtie \underline{X} =_{MVD} \bowtie \underline{Y}$.
Then $\bowtie I(B) =_{MVD} \bowtie I(B')$. Therefore by Corollary 5, $B = B'$. Now the equality $\underline{X} :=: \underline{Y}$ follows from ($\underline{X} :=: I(B)$ and $I(B) :=: \underline{Y}$) by transitivity (Corollary 6). □

Because of the uniqueness of $B \subseteq P(U)_{-2}$ such that $j =_{MVD} \bowtie I(B)$ (Corollary 5), we can summarizingly state as follows.

Theorem 7 (conclusion)

1. Every equivalent class $[\underline{X}]$ which is an element of the quotient set (Lemma 5). $C(U)/ :=:$, can be uniquely represented by $[I(B)]$ with $B \subseteq P(U)_{-2}$.
2. Such a B is effectively computable from any element $\underline{X}$ of the class (emma 5).
3. Moreover, $\forall \underline{Y} \in C(U)$ $(\underline{Y} \in [I(B)] \iff I(B) := \underline{Y})$.
4. The subset of JD(U) defined by $[\bowtie I(B)] = \{\bowtie \underline{X} \mid \underline{X} \in [I(B)]\}$

is an equivalence class of JD(U), which is an element of the quotient set JD(U) / $=_{MVD}$.

5. Let M be a subset of MVD(U) and let B be a subset of $P(U)_{-2}$

   $[\bowtie I(B)] = jds(M) \Longleftrightarrow B$ is a basis of M. □

The correlations are shown in the following figure.

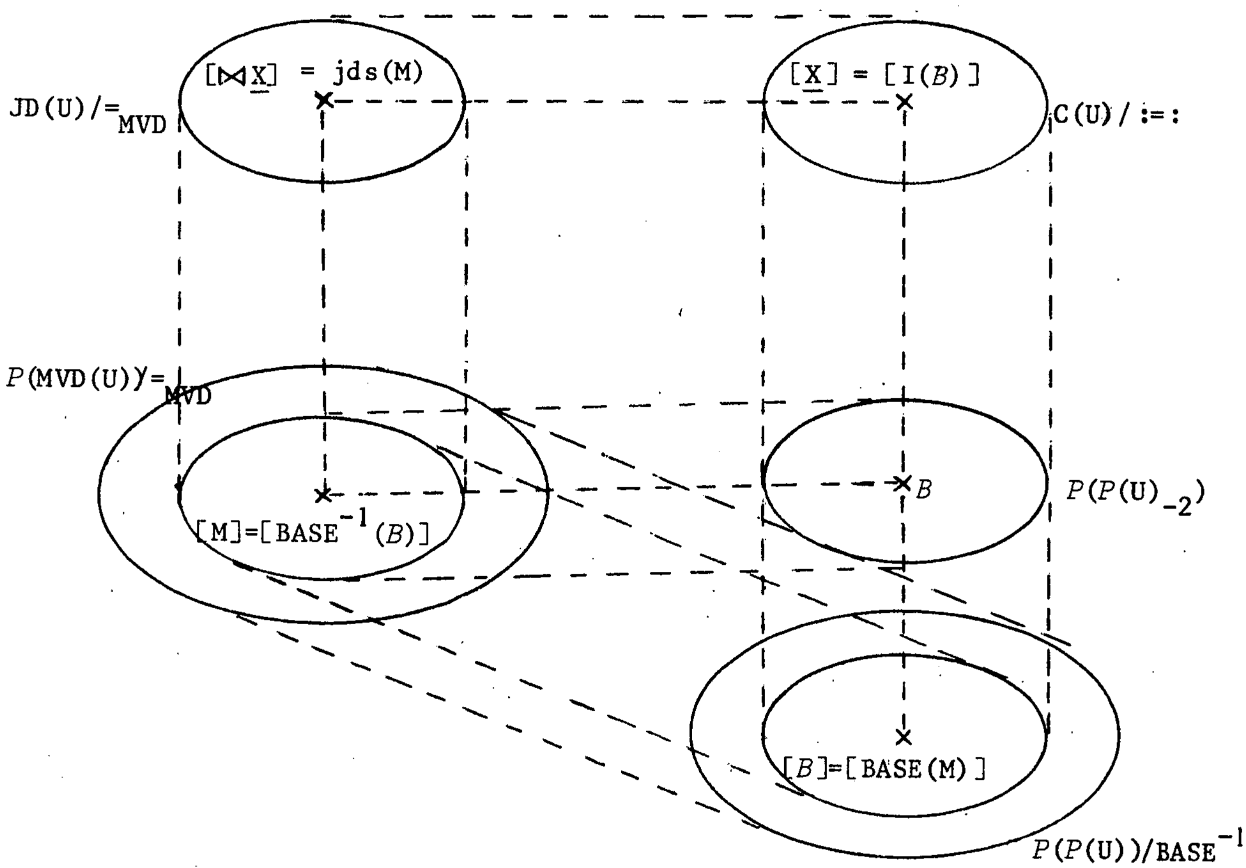


Figure 1

EXAMPLES

Throughout the following examples, we consider the universe of attributes $U = \{x,y,z,v,w\}$. Let us denote the elements of $P(U)$ by x, xy etc. instead of $\{x\}$, $\{x,y\}$ etc.

1. Consider the set of MVD's $M = \{z \twoheadrightarrow v,\ v \twoheadrightarrow w\}$. Clearly,

   $$\mathrm{BASE}(M) = \{\emptyset\} \cup \{x,y,w\} \cup \{xy,xw,yw,vw\} \cup \{xvw,yvw,zvw,xyw\} \cup P(U)_{-1} \cup \{U\},$$

   where $P(U)_{-1} =_{\mathrm{def}} \{X \mid X \in P(U) \text{ and card } X = \text{card } U - 1\}$.

   BASE(M) is naturally a basis of M. But $\mathrm{BASE}(M)_{-2}$ is not a basis of M, because :

   $\mathrm{BASE}(M)_{-2} \supset\!\!\sim x \twoheadrightarrow w$,

   $\mathrm{BASE}(M) \not\supset\!\!\sim x \twoheadrightarrow w$.

   That is, $\mathrm{BASE}^{-1}(\mathrm{BASE}(M)_{-2}) \not\subseteq \mathrm{BASE}^{-1}(\mathrm{BASE}(M))$. So by Theorem 3, we can conclude that there is no JD on U such that $j =_{\mathrm{MVD}} M$.

2. Let $B = \mathrm{BASE}(M)_{-2} \cup \{xy\}$, with $\mathrm{BASE}(M)_{-2}$ as above. For this B we can verify that $\mathrm{BASE}^{-1}(B) = \mathrm{BASE}^{-1}(\mathrm{BASE}(M))$. That is to say, B is a basis of M, by definition. Naturally, $\mathrm{BASE}^{-1}(B_{-2}) \not\subseteq \mathrm{BASE}^{-1}(B)$.

3. Now, take $M' = \{z \twoheadrightarrow v,\ \emptyset \twoheadrightarrow w\}$. Then $\mathrm{BASE}(M') = \{\emptyset\} \cup \{w\} \cup \{xw,yw,vw\} \cup \{xvw,yvw,zvw,xyw\} \cup P(U)_{-1} \cup \{U\}$.

   This time, we can prove that

   $\mathrm{BASE}^{-1}(\mathrm{BASE}(M')_{-2}) \subseteq \mathrm{BASE}^{-1}(\mathrm{BASE}(M'))$.

   As BASE^{-1} is an order-inverting function, the inverse inclusion holds. That is to say, $\mathrm{BASE}(M')_{-2}$ is a basis of M'. So, by Theorem 7,

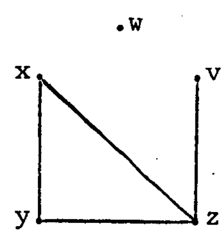
   $\mathrm{jds}(M') = [\bowtie I(B')]$

with B' = {xvw, yvw, zvw, xyw} (=BASE$(M')_{-2}$).

By definition,

$$I(B') = \{yz, xz, xy, zv, w\}.$$

The set $I(B')$ is shown pictorially in Figure 2.



$I(B')$

$I(B')$ contains the set Clique (xyz) = {yz, xz, xy}. So, by the reduction rule (r1),

$$I(B') :=_1 I(B') \cup \{xyz\}.$$

Now, $I(B') \cup \{xyz\}$ allows us to apply the reduction rule (r2) on it. Thus we have :

$$I(B') := \{xyz, zv, w\}.$$

The only deduction rule applicable to {xyz, zv, w}, is the rule (r3). The only elements of $[I(B')]$ are $I(B')$, {xyz, zv, w} and those obtainable from these two by successive application of the deduction rule (r3). The elements of jds(M') are of the form $\bowtie \underline{X}$ with $\underline{X} \in [I(B')]$.

4. Let $\underline{X} = \{xyz, yzv, xzv, vw, xw\}$ and let $\underline{Y} = \{xyz, yzv, xyv, xvw\}$. Then Clique $\underline{X}$ = Clique $\underline{Y} = \{xy, xz, yz, yv, zv, xv, vw, xw\}$. Here Clique B, with $B \subseteq P(U)$, is a subset of $P(U)$, defined as follows :

$$\text{Clique } B =_{\text{def}} (\bigcup \{\text{Clique } (B) \mid B \in B \text{ and card } B \geq 3\}) \bigcup$$

$$\{B \mid B \in B \text{ and card } B \leq 2\}.$$

So by Lemma 3, we can conclude that $\underline{X} :=: \underline{Y}$. In fact, let $\underline{Z} = \{xyz, yzv, xzv, xyv, xvw\}$. Then $\underline{X} :=: \underline{Z}$ and $\underline{Y} :=: \underline{Z}$.

Let us denote by $\underline{W}$ the set which is equel to Clique $\underline{X}$ (and therefore to Clique $\underline{Y}$). The graph of $\underline{W}$ is shown in Figure 3, and we see that

$\underline{W}$ = Clique(xvw) $\bigcup$ Clique(xyzv).

We can therefore conclude that

$\underline{W} :\doteq \{xvw, xyzv\}$.

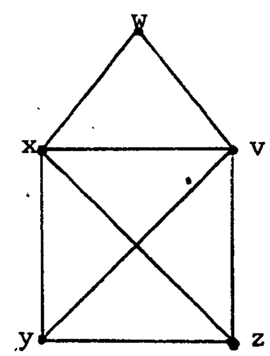


Figure 3

We can also verify that $\underline{Z} := \{xvw, xyzv\}$, with intermediate steps, $\underline{Z} \bigcup$ Clique (xyzv) and $\underline{Z} \bigcup$ Clique (xyzv) $\bigcup \{xyzv\}$.

Now let us note that $\underline{W}$ = Clique(xyzvw) $- \{yw, zw\}$.

So the subset B of $P(U)_{-2}$ with the property : $I(B) = \underline{W}$, is the following :

$B = P(U)_{-2} - \{xzv, xyv\}$.

Clearly $M = \{xv \twoheadrightarrow w\}$ is a subset of MVD(U) such that B is a basis of M. So, $\bowtie \underline{X}$, $\bowtie \underline{Y}$, $\bowtie \underline{Z}$, $\bowtie \underline{W}$ are all MVD-equivalent to a single MVD $xv \twoheadrightarrow w$.

5. Suppose that an MVD $d = X \twoheadrightarrow Y$ is given and that we want to know the subset B of $P(U)_{-2}$ which is a basis of d. To do this, we consider the covering $\{X \cup Y, X \cup (U-Y)\}$, from which we compute $\underline{W}$ = Clique $(X \cup Y) \cup$ Clique $(X \cup (U-Y))$. Then $\underline{W}$ must be of the form $I(B)$, from which we can compute $B \subseteq P(U)_{-2}$ that we are looking for.

ACKNOWLEDGEMENT

I am indebted to Nicolas SPYRATOS for his careful reading of this paper.

BIBLIOGRAPHY

[A] ARMSTRONG, W.W.
"Dependency structures of data base relationships"
in INFORMATION PROCESSING 74, North-Holland Publ. CO. 1974.

[BFMY] BERRI, C., FAGIN, R., MAIER, D., YANNAKAKIS, M.,
"On the desirability of acyclic database schemes"
Research Report IBM, RJ 3131, 1981.

[FMU] FAGIN, R., MENDELZON, A.O., ULLMAN, J.D.
"A simplified universal relation assumption and its properties"
Research Report IBM, RJ 2900, 1980.

[H] HANATANI, Y.
"Bases for FD-MVD structures"

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique